The University of Bradford Institutional Repository

Link to original published version: http://dx.doi.org/10.1108/IJOPM-02-2015-0084

Citation: Matthias O, Fouweather I, Gregory I and Vernon A (2017) Making sense of big data – Can it transform operations management? International Journal of Operations and Production Management. 37(1): 37-55.

Making Sense of Big Data – Can it Transform Operations Management?

Structured Abstract:

Purpose
This paper focuses on the application and exploitation of Big Data to create competitive advantage. It presents a framework of application areas and how they help the understanding of targeting and scoping specific areas for sustainable improvement. Empirical evidence demonstrates the application of Big Data in practice and tests the framework.

Design/methodology/approach
An exploratory approach is adopted to the secondary research which examines vendors' offerings. The empirical research used the case study method.

Findings
The findings indicate that there is opportunity to create sustainable competitive advantage through the application of big data. However there are social, technological and human consequences that are only now beginning to emerge which need to be addressed if true long-term advantage is to be achieved.

Research Limitations
The research develops a framework and tests it only in 2 dimensions. This should be expanded. The vendor analysis limitations lie within the nature of the information available and the difficulties in mitigating against bias.

Practical Implications
The suggested framework can help academics and managers to identify areas of opportunity to do so, setting new levels of performance and new agendas for business.

Originality/value
This work contributes to service operations management, building on Kranzberg (1986) and the impact of technology and on Fosso Wamba *et al.* (2015) by developing a systems application framework to further understanding of big data from a practical perspective to extend their research taxonomy insights. Our case studies demonstrate how the use of BD enhances operational performance.

Keywords
Big Data, Operations Management, Data Analysis, Business Performance

Making Sense of Big Data – Can it Transform Operations Management?

1. Introduction

Technology is a major source of change in today's business world. As internet and mobile technologies grow, all aspects of our lives are being transformed. Enterprises generated and stored an estimated seven exabytes (billions of gigabytes) of data in 2010, whilst consumers generated an additional six exabytes (Manyika *et al.*, 2011). The IDC (2012) estimates that by 2020 this will increase forty. Exponential technological growth can bring social and organisational challenges. As Kranzberg (1986:545) notes:

"*Technology is neither good nor bad; nor is it neutral*… *technical developments frequently have environmental, social and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves*".

Technology companies (typified by Google, Microsoft, IBM, Oracle and SAP) and commentators such as Bell (2013) and Porter and Heppelmann (2014) promote the need for businesses to exploit datastreams to create competitive advantage. The OECD (2013:4) suggested that:

"The exploitation of data promises to create added value in a variety of operations ranging from optimising the value chain and manufacturing production to more efficient use of labour and better customer relationships".

If organisations can harness internally or externally generated data, their operational capabilities could be transformed. Whether by providing greater volumes of reliable and timely information for decision makers, or by automating decision-making processes, there is a collective assumption that Big Data (BD) will benefit organisations, individuals and society (Bughin *et al.*, 2011, Barton and Court, 2012). Computer applications capable of analysing huge data-sets are becoming readily available (Fisher *et al.*, 2012). Opresnik and Taisch (2015) highlight that the challenge for organisations is to develop strategies that exploit BD to generate added-value. Others question whether BD will ultimately deliver what it promises (Croxall, 2014, Lury, 2013). Gandomi and Haider (2015) observe that there has been little critical discourse, or empirical academic research, into BD and how it might be harnessed.

The research described in this paper arose from a curiosity about the possible benefits of BD and the lack of clarity about exactly how amassed data can be helpful to organisations. The paper empirically examines current application of BD to create competitive advantage, using current literature is examined and secondary research to evaluate applications. This broad assessment of systems solutions purchasing enables the development of a framework indicating applications usage. The framework is then evaluated using as case studies two UK retail organisations which sought to exploit available BD.

The paper adds to a small but growing body of Operations Management literature, alongside work such as Demirkan and Delen (2013), Fosso Wamba *et al*.(2015), and Huang and Handfield (2015). It makes a contribution to knowledge by developing a framework detailing how BD applications are currently being exploited. It contributes to practice by helping organisations develop their operational strategy to best exploit the available data to engender sustainable improvement.

2. **Literature Review**

There is considerable hype around the term BD (Gandomi and Haider, 2015, George *et al.*, 2014, Gartner, 2013, 2015). It has become endemic since the emergence of the term in the mid-1990s (Diebold, 2012). Technology companies, whether providers of ERP systems, CRM software or Business Analytics have been promoting BD for over a decade. In 2012 McAfee and Bryynjolfsson, writing in the Harvard Business Review (2012), said BD represented a '*revolution in management'*. Brown *et al.* (2011) asserted that it would "*transform business processes and alter corporate ecosystems"*. Manyika *et al.* (2011) described it as *'the next frontier for innovation, competition and productivity'*.

The rhetoric of these messages seems to be shaping the expectations of academic communities. Fosso Wamba *et al*. (2015) provided a comprehensive review of existent literature on BD. They concluded that the majority of publications focused on BD technologies and access to data. To broaden the understanding of the role BD has in capturing business value they developed a general taxonomy from this review. Empirical research by Wieland *et al.* (2014) suggests that within the field of Supply Chain Management academics expect BD to be 'the hot topic' for the next five years. They indicate that leading researchers acknowledged a degree of scepticism that it could be the latest in a long line of management fashions and fads. This view is supported by Madsen and Stenheim (2013).

2.1 The Antecedents of Big Data

The ability to generate vast streams of data has increased as the ability to rapidly process that data has increased. Moore's Law stated that the processing power of computers doubled every eighteen months (Moore, 1965). Kryder's Law says that digital storage is increasing at a similar rate to data processing power (Esener *et al.*, 1999). Hruska questions whether Moore's Law remains valid and suggests that current storage technologies are reaching their limit (2013).

In the 1990s systems engineers acknowledged that greater volumes of data, generated at higher rates, do not automatically lead to more information and knowledge. Without the ability to analyse and understand the data being generated, possessing more data may actually be a hindrance. Unless data can be transformed into information that facilitates good decision-making and enhances operational performance, it has no purpose (Jifa and Lingling, 2014). Unlocking the knowledge within the data remains the central challenge.

The link between data and information has been explored by systems theorists for decades and the insights provided have shaped operations management for many years (Carvajal, 1992). It is through the "*application of data and information*", that knowledge is created (Ackoff, 1989). Knowledge is the ability to use information within a particular context. Sherdoff's (1994) DIKW hierarchy links Data, Information, Knowledge and Wisdom. It provides a theoretical pyramid built on a foundation of data with each successive layer resting on the one beneath. Focus has generally been on the three lower levels of the hierarchy (Rowley, 2007).

Information is:

"organised or structured data, which has been processed in such a way that the information has relevance for a specific purpose or context, and is therefore meaningful, valuable, useful and relevant" (Rowley, 2007).

Put more practically, by processing data appropriately it becomes "*useful for decisions and or action*" (Liew, 2007). If information is "know what", knowledge is "know why". Knowledge permits decision-making, enabling the selection of a particular action from a range of possibilities and leads to the "know how" to improve operational performance and create competitive advantage. With bigger (and by implication better) data comes the promise of enhanced knowledge and decision-making.

2.2 The Emergence of Big Data

There is no universally accepted definition of BD. McKinsey Global Institute's study (2011) suggests that: *"… data-sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse."* The Gartner definition is " *Data assets that require innovative forms of information processing for enhanced insight and decision-making*" (2014). Both seem to imply that the challenge of BD is scale, or volume, omitting any mention of the human aspect, or the need to understand how meaning is created from BD, despite Kranzberg (1986) identifying this issue some 30 years earlier.

Although the origins of the term BD are obscure, there is agreement about when it appeared in academic literature (O'Leary, 2013, Jifa and Lingling, 2014). Cox and Ellsworth presented a paper at the IEEE's 8th conference on Visualization (1997). Their narrative presented BD as a problem for systems engineers pondering how large data-sets that exceeded available memory capacity could be managed. Subsequently Laney (2001) published a research paper in which he predicted that with the rise of e-commerce, enterprises would need to manage ever larger data-sets and the challenge would be in terms of three dimensions: Volume, Velocity and Variety. The 3Vs have become central to understanding big data. Essentially BD refers to high velocity, variable and complex data requiring advanced technologies and techniques to capture, store, distribute and manage for subsequently useful analysis. Over time the 3Vs have been extended with the addition of additional dimensions: Variability, Veracity and Value. However, the 3Vs remain core for any understanding. There has also been a shift away from considering the challenges of BD to a more speculative approach that looks towards the future benefits of BD.

Traditional business analytics have focussed on internally created data. An essential element of the discourse on BD is social data (Chae, 2015). The proliferation of smart technologies has created a constant stream of data about individuals that can be utilised. Additionally, social media offers a real-time window into people's opinions, wants and needs, not previously accessible. Within Operations Management the focus has been on internal data related directly to operations and processes whilst considering the behaviour and expectations of customers. BD offers new ways to understand the external environment (Dubey *et al.*, 2015).

A complication in utilising social media-generated data which has had little attention is how customers are affected. With data and information at their fingertips, customer expectations regarding the service they receive change. Customers create their own choices and their own relationship with a business by the online searches they make. They expect interaction through a channel of their choice at a time convenient to them with the same experience regardless of channel or device used. Inevitably changes in managing customer-facing processes. Most importantly perhaps, the dynamics of trust and co-production are altered. Little research exists on this in mainstream Operations Management.

2.3 Big Data Research

Gandomi and Haider (2015) and Fosso Wamba *et al's.* (2015) work shows the rapid growth in the use of the term Big Data in publications and corroborates Wieland *et al.'s* (2014) description of it as a “hot topic”. As more articles emerge, the significance of the term increases and a positive feedback loop creates greater usage, further amplifying the significance collectively attached to the term, as illustrated by the *twitter storms* phenomenon (Segerberg and Bennett, 2011).

Yet research on data analytics and operations improvement is not new. For instance, Bell (1999:307) cites a FedEx executive referring to the RFID transmitter, indicating his organisation had “*succeeded by applying scientific methods to its operations*” utilising advanced analytics. The RFID transmitter is a smart object still referred to in mainstream operations management papers, and now in BD (Ilic *et al.*, 2010, Zelbst *et al.*, 2012, Meyer *et al.*, 2014, Lee and Özer, 2007, Zhong *et al.*, 2015). In their recent work (2015) Chongwatpol and Chan describe a case study which used a large dynamic data-set to enhance operational decision-making and increase effectiveness. Their research shows how data analytics can be used to find alternative ways of assessing business issues. They did not use the term BD.

Thus, whilst empirical BD research is not yet widespread in the operations management community, the technologies and analytic techniques required to deliver BDs promised value are well-researched.

2.4 The Big Idea of Big Data

In 2013 almost one billion smartphones were sold (Gartner, 2014) each capable of creating and collecting masses of data. In an operational environment smart machines are commonplace; new equipment comes with an array of sensors and data trackers that produce and store endless data (Lucke *et al.*, 2008, Meyer *et al.*, 2014, Zelbst *et al.*, 2012). Technology companies are keen to develop more applications that will generate even more data. Customer transactions and electronic feedback can be processed to capture demand and levels of satisfaction. Internally myriad smart technologies generate instantaneous feedback on the status and performance of internal resources. The claim is that processes can be controlled more effectively and better decisions to exploit opportunities and solve problems taken because of the information generated by new technology.

The Google Flu Trends Project (GFT) reveals just how such claims can prove disappointing (Lazer *et al.*, 2014). The project’s goal was to develop a means of identifying the emergence of flu so that health resources could be mobilised to treat the illness and prevent an epidemic. Using readily available data from Google’s search engine, data on the frequency of “flu” searches was collected. A sophisticated analytical tool was developed to extrapolate from the search data to predict future rates of flu. After apparent early success the model was found to consistently over-inflate future occurrences and was less accurate than existing ‘small data’ strategies that utilised data on confirmed cases of flu. Their work highlights how the absence of criticality and rigour leads to flawed results: “*quantity of data does not mean that one can ignore foundational issues of measurement and construct validity and reliability and dependencies among data*” (ibid).

BD usefulness is constrained by the ability of the researcher to ask the right question, the same constraint as in 1943 when Abraham Wald was analysing appropriate positioning of armour on fighting aircraft. The predominant thinking was to armour those areas where it was obvious that aircraft had sustained damage. Wald’s wondered where the damage might be in those aircraft that had not returned. The only data he had was the analysis of bullet holes on planes that had returned. He successfully developed an assessment methodology used in subsequent wars (notably Korea and Vietnam) to minimise enemy-inflicted damage (1980).

In the case of the conclusions drawn from the GFT data, the absence of criticality, curiosity and validity resulted in erroneous projections. There is no way of knowing who searches for data on flu, or, equally importantly, why. The data collection only demonstrates the population that has an interest in flu - people with symptoms, and/or their friends, family, work colleagues. Data mining and analytic techniques may reveal interesting patterns hidden within large datasets. Without an understanding of what the patterns reveal, the information may have little value or, worse, lead to incorrect conclusions. As GFT shows, searching for patterns is complex. Lazer *et al.’s* message is clear: it is the insight and understanding of those involved in the analysis of data (whether big or small), that is central in creating value. The example of Wald remains an important reminder.

Uncritical analysis of poorly understood data-sets does not generate knowledge. As Li *et al.* (2015:3) acknowledge, despite the increasing availability of data, how BD can be used to support decision-making is “an enormous challenge”. Whatever the size of the data-set, it needs appropriate analysis to create useful information that reveals what is significant within the data. “*Not all information is useful for improving our understanding and judgements*” (Saaty, 2008) and too much information can create uncertainty, hindering decision-making. A certain amount of inference is required, and possibly the use of Bayesian analysis in support. Only then will the available data facilitate effective decision-making that can add value to operations and enhance performance.

Piccoli and Pigni’s (2013) work illustrates how Digital Data Streams (DDS), can be used within an operational setting to *replace* routine decision-making activities with predetermined (mindless) responses. This proven technology creates far higher levels of automation than previously possible, but the authors stress that DDS is not BD. Rather, it is an example of closed-loop feedback control. Research by Saetre *et al.* (2003) found that over-reliance on technology can create mindless responses if individuals fail to contextualize information, more likely to reduce operational performance than improve it. If BD is to be exploited to assist in decision-making it needs to be analysed and utilised with care.

Bisel *et al.* (2014) indicate that there is an assumption that the size of a data-set is seen by many researchers and analysts as a proxy for quality, and yet as Lazer *et al.* (2014) stress, quantity is not a substitute for quality. The volume and variety of BD make assessing its veracity challenging. Whilst statistical tools can be used to filter out erroneous and missing data, establishing the veracity of a large data-set is not straightforward. Chongwatpol and Chan (2015) illustrate just how much work is required to ensure veracity of data and the information generated from large dynamic datasets.

Significantly, this is not a matter of technology, but one of skills. What new skills are required is something that writers do not necessarily agree upon (Miller, 2014). The OECD (2013:29) acknowledges that an appropriate mix of advanced ICT, statistics and specific sector skills are required. Fawcett and Waller (2014) similarly recognise the need for both data and domain skills, but stress that it is the ability to apply technical skills that is important. Manyika *et al.* (2011) highlighted that organisations do not have the talent to derive insights from BD. When presented with a complex and ever-changing stream of data and information, the ability to think creatively, grasp the situation and act accordingly may require a diverse range of attributes, such as intelligence, intuition, imagination and creativity, not always recognised or valued in the workplace. Despite advances in AI, technology cannot provide these attributes. Mainstream media (Philipson, 2014) recently reported that the British Intelligence Service (GCHQ) recognised this and employed more than 100 dyslexic and dyspraxic analysts because of their skills in identifying patterns and ability to analyse complex data.

Analysing BD is complex for a variety of reasons. Traditionally data analysis has considered discrete data that can be handled using well-established and sophisticated quantitative techniques, such as data feeds from sensors. Processing such data is straightforward, and can easily be automated. The data generated through social media presents a greater challenge. It is unstructured and comes in a range of formats, often with multimedia content and threads of previous textual dialogues embedded within it. As Bisel *et al.* (2014) note, analysing such qualitative textual data requires specific skills. Analysts are working on tools to achieve this, but the ability of software to analyse text remains rather limited (Chen *et al.*, 2012). One exception is the work of Bollen *et al.* (2011). They successfully analysed the 'mood' of Twitter feeds to predict movement in the Dow Jones Industrial Average, identifying an accuracy of 86.7%. Yet Tinati *et al.* (2014) observe that the current forms of analysis of rich unstructured qualitative data is limited to classifying, linking, and revealing distributions of words, reducing it to little more than a word-count survey. Whilst these are essential processes of data reduction which allow analysts to sift through large data-sets quickly, they leave much unexplored. The role of the analyst is central to the process of interpreting the data, as demonstrated by Wald. Housely *et al.* (2014) detail work that is being undertaken to enhance capabilities in this area, but sophisticated tools and techniques have yet to be established.

There is an additional dimension to complexity that is often overlooked when considering the challenges of BD. Jifa and Lingling (2014) highlight the problems of working with OCG (open, complex, giant) systems. Their insight reveals that the internet creates an interconnected system, and the data being harnessed is not simply more data about an existing system, it is further complicating the system. Nor is it a closed system of the sort created within operations; it is a system constantly changing, with positive and negative feedback loops interacting dynamically, which can thereby simultaneously create stability and chaos. This confirms GCHQ's conclusion that using data from open complex systems requires a different set of skills. These are the sort of skills dyslexic pattern-spotters have been forced to develop to survive in a text-orientated world. Their ability to spot oddities in patterns, which must have been useful survival skills for hunter-gatherers, are now being applied to data-sets rather than animal behaviours, and the outcome is insight into behaviour.

2.5 Literature Summary

BD describes the large quantity of data generated and stored by modern technologies which has forced businesses to consider how they exploit the resultant information flows. Collecting data from multiple channels has never been easier. The idea that data shadows and information trails of people, machines, commodities and even nature can reveal secrets because the power and prowess now exists to uncover them is alluring. The hype says that the more data gathered, the better the decision-making. Reality is less straightforward. Usability is critical. Data needs to be converted in a robust and reliable way to be translated into knowledge to be applied. Agility and flexibility in data collection is good, but it is necessary to connect and correlate relationships, hierarchies and multiple data linkages, otherwise the data remains meaningless. Meaningless data is the digital information equivalent of a rubbish heap, as shown by the GFT example.

In an ideal world, big data would help organisations set new levels of performance and new agendas for business. This research explores how advances are being made that make this a possibility and considers the challenges that remain.

3. **Methodology**

The research objective was to understand if organisations use BD and if its benefits match its widely presented potential. Given that this is a relatively new research area and the research largely exploratory, to achieve the objectives a two-phase multiple case study approach was adopted (McCutcheon and Meridith, 1993). In the first phase secondary data from BD vendors informed the development of a theoretical framework categorising current BD usage. Phase two explored two organisations that have sought to exploit BD to evaluate whether the framework is supported by empirical data.

3.1 Phase One

Products that are available influence how organisations are likely to exploit BD. An exploration of the types of application that BD vendors offer brings insights into the ways that BD is being used. A purposeful sampling strategy was adopted because it was sympathetic to the objectives and the exploratory nature of the research. An earlier study had already identified the top ten vendors of BD solutions (Kelly, 2014). This was validated against other sources of data (industry and market reports, trade journals and media coverage) and judged to remain valid for the purposes of this research and was therefore the sample used.

Individual case studies produced by the vendor and made available on their website that related only to applications of BD were examined. To mitigate against the risk of obsessive positivity, only case studies with a named client were used, based on the premise that the client would not allow their reputation to be damaged by exaggerated vendor claims. A paragraph describing each case was created.

Cases relating solely to the implementation of technology infrastructure were rejected, as this is independent of the application. Three of the vendors did not provide case studies or provided them in an anonymous form, restricting the vendor sample to seven. Within this sample, some vendors provided numerous case studies and others few. A total of 253 case studies were identified across a wide range of industrial sectors. There was deliberately no attempt to establish in-case validity or consistency through triangulation. Rather, this phase of the research sought to identify the range of product offerings, not to classify the vendors' products. This may have led to an unrepresentative data-set, but establishing the range of applications was the objective at this stage.

Eisenhardt (2002) indicates that between 4 and 10 case studies is an ideal number to develop a theory, so numerically the validity of the approach adopted is high. Grounded theory is used to build a framework of BD usage based on the client track record of the leading BD vendors, in keeping with Glaser and Strauss (1968).

With the number of usable case studies amongst the vendors varying widely, it was necessary to mitigate bias towards any particular vendor based on proclivity to publish cases. Client track record is primarily a mechanism for marketing the services of a professional services firm or technology vendor. The case studies that comprise their track record are therefore universally positive and likely to be skewed towards applications that are more profitable. Consequently, a limitation of the methodology is that unprofitable implementations of BD may be missing from any analysis undertaken. This is an area for further research.

3.2 Phase Two

The second phase of the research examined the experiences of two UK retail organisations as they embarked on performance improvement projects using BD. In this instance the sampling was opportunistic. Readily available data and direct access rather than reliance on self-reporting were more important factors than an alternative sampling strategy.

The case studies, their drivers, method and outcome, are outlined separately. They are compared with the general literature on BD, what it is and what it is supposed to do (Wieland *et al.*, 2014, Fosso Wamba *et al.*, 2015, Haas *et al.*, 2014). The cases are applied to the vendor analysis framework developed from the earlier phase. The aim was to ascertain how organisations currently leverage BD and if there is anything to differentiate this from standard data analysis techniques.

Aligning the two phases, the research sought to compare the hype with the reality in an operational setting. Thus the operational implications of BD usage for sustainable competitiveness are highlighted whilst considering the social, technological and human consequences, in line with Kranzberg (1986). As such the research complements the work undertaken by Dubey *et al.* (2015). Both adopt a two-phase, mixed-methods, sequential approach of theory-building, followed by the evaluation of primary data to provide a degree of theory validation.

4. **Framework Development: The Applications of Big Data**

Content analysis was undertaken for each of the descriptive paragraphs of the case studies, and a list of categories developed. Each case study was categorised with one or more types. Some had several threads, so were treated as more than one content item. Whilst some categories were far more prevalent than others in the published cases, all applications, even if there was only a single example, were incorporated into the analysis. Hence, by controlling for the popularity of an application, the research covered all applications equitably. A set of aggregate classifications was developed and again each paragraph was classified.

The categorisation identified a number of application areas which resulted in the development of a framework on two axes: scale and time horizon, shown in Table 1.

Table 1: Typology of BD Application

Scale / Time Horizon	Micro (1)	Macro (2)
Future	Prediction	Service Design Strategic Support
Present	Personalisation Detection	Optimisation
Retrospective	Troubleshooting	Compliance

The scale axis considered whether the case study provided BD services that focused on macro applications, e.g. analysing flows within a supply chain, or micro applications, e.g. providing recommendations to a single customer. The time horizon scale looked at whether the focus of the data was regarding the past, for example historical data analysis to demonstrate compliance, the present, for example real-time fraud detection, or modelling the future, for example new service proposition design. Table 2 shows the full description for all the categories. The framework was then validated against the stated service offerings of the vendors to ensure that all of their offerings could be classified by within the typology. This was done by reviewing the section of vendor websites relating to services provided and plotting each service back onto the framework.

Table 2: Description of typological categories

Category	Description
Prediction	Using BD to predict what products an individual customer would want in future, for example the promotion of recommended purchases
Service Design	Using the data feeds from existing products and services to inform the design of new product or service propositions
Strategic Support	Using BD to inform strategic decisions
Personalisation	Using BD to personalise the experience of an individual customer, for example the provision of information that they were likely to want
Detection	The use of BD to identify issues in real-time, for example fraud detection
Optimisation	The use of BD to optimise performance in a process, for example through its use in supply chain planning
Troubleshooting	Use of BD to identify the causes of problems that have occurred
Compliance	Use of BD to demonstrate compliance to regulations within an organisation. This includes using BD to analyse complex sources of data such as recordings of telephone conversations

This typology provides a framework for the classification of the application of BD in Operations Management and the wider field of management. The next step was to test its validity by examining two case studies.

5. The Case Studies built on Primary Data

5.1 Case Study One – Online White Goods Retailer

Background

This organisation wanted to deliver tailored experiences for new online customers when they visited the company's website. The absence of physical clues about a customer such as gender, clothing and ethnicity that are used in face-to-face encounters meant that the statistically significant real-time sales discrimination based on appearance was not possible (Wise, 1974). An employee was musing one day and wondered if, in the absence of physical clues, the technology the customer used could help the company understand the customer better. The research question was: "Can historical access data deliver more tailored experiences to new customers when they visit the website?"

Data Collection

Search advertising is the primary tool directing potential customers to the website and is complemented by display advertising where defined criteria about context and previous browsing behaviour are met. Conversion rate optimisation (CRO) is used to increase the number of website visitors booking a design visit and thereby moving more successfully through the company's sales funnel. CRO positively influences cost-per-lead and permits analysis of the technology in the form of operating system (OS) and browser customers use to access the website. OS and browser data combined was collected for a 12 month period (2014). 1,621,262 website clickstream data was pulled into the company's CRM system at the point of design visit enquiry, and analysed.

Data Analysis

Using Google Analytics to investigate internet traffic data it became apparent that the research question could be more precisely defined:

1. Can customer technology use for website navigation predict purchase type and spend?
2. Can this information be used to deliver a more tailored website experience?

Preliminary analysis showed that Microsoft Explorer (IE) and Firefox usage declined by 40% in 12 months while all other browser traffic increased. Devices were identified as the most important influencer because most users stay with the default browser provided (Browser-update.org, 2015). As use of smartphones and tablets increased, so did access to the company website using Safari and Google Chrome, with corresponding decreases in IE and Firefox OS.

Three separate hypotheses were developed to test the correlation between purchasing choice and variables associated with web use that could be captured using Google Analytics.

H1 – Product style preference and device are associated

H2 – Product style preference and operating system are associated

H3 – Product style preference and browser are associated

The clickstream data was segmented into 3 technology categories and number of visits to the top 4 styles of the 2 product categories the company offered. Pearson's Chi-square test of association was selected to test the hypotheses because of its versatility and ability to deal with categorical data (Hair *et al.*, 2007). Statistical significance was tested to 95%, judged by the organisation to be an appropriate level of confidence to use in the tests.

Table 3: Case Study One. Hypotheses Testing

H1 – There is association between product preference and the device used to browse	
H_0	Product preference and device are not associated
H_1	Product preference and device are associated
Significance Level	$\alpha = 0.05$
Pearson's Chi Square Test	$(\chi^2) = 509.2441896$
Test Statistics	$P = 2.6244 \times 10^{-111}$ Critical Test = 5.991464547
Evaluation	$(\chi^2) >$ Critical Test $p > \alpha$ **From both tests Reject H_0 and accept H_1**
H2 There is an association between product preference and the operating system used whilst browsing.	
H_0	Product preference and operating system are not associated
H_1	Product preference and operating system are associated
Significance Level	$\alpha = 0.05$
Pearson's Chi Square Test	$(\chi^2) = 839.9343174$
Test Statistics	$P = 4.4630 \times 10^{-177}$ Critical Test 14.06714045
Evaluation	$(\chi^2) >$ Critical Test $p > \alpha$ **From both tests Reject H_0 and accept H_1**
H3 There is an association between product preference and the browser used	
H_0	Product preference and browser are not associated
H_1	Product preference and browser are associated
Significance Level	$\alpha = 0.05$
Pearson's Chi Square Test	$(\chi^2) = 685.4852838$
Test Statistics	$P = 6\ 7527 \times 10^{-146}$ Critical Test 11.07049769
Evaluation	$(\chi^2) >$ Critical Test $p > \alpha$ **From both tests Reject H_0 and accept H_1**

Thus it was possible to accept all three hypotheses.

Because of the interdependency between the operating system and browser, a fourth hypothesis was developed:

H4 – Product style preference and [OS + browser] are associated

A total of 21 different combinations of operating system and browser were identified from the data available. Again using Pearson's Chi square test and the same significance level as before, a p value of 6.1098×10^{-195} indicated this hypothesis could also be accepted.

Similarly the critical value of 31.4104 was significantly below the test statistic (χ^2) of 980.3270 .

Findings

The result from H4 definitively answers research question 1and confirms that assessing customer preference based on technology is viable. Combining OS and browser variables provided a rich set of data segments to use in tailoring the customer experience. Mac OS and Google Chrome suggested a preference for a modern, expensive products whilst IE meant more traditional style and careful spend. Research question 2 was also answered from test 4. The strong association between browser/OS combinations and product style preferences enabled the business to identify three customer categories based on the technologies used for browsing:

- more interested in traditional
- more interested in modern
- no bias

The website was redeveloped to act upon insights produced from the statistical analysis of browsing behaviour data. By collecting information about the OS and browser combination during the initial online contact and applying the knowledge of likely customer preferences, customers could be direct towards specific webpages tailored to meet their likely preferences. Customers benefited by quicker navigation to the products they were likely to be interested in, with a corresponding improvement in the “hits:design visit” ratio.

Considering the case carefully, it is clear that the analysis undertaken was relatively straightforward, but the improvement in sales would not have been possible without the right question being asked and an association being found between hitherto unmade combinations.

When assessed according to the framework presented in section 4, this retailer’s approach falls into the category of *Personalisation*.

5.2 Case Study Two – Multi-Channel Retailer

This company had a vision: to create a competitive edge by being the most trusted UK provider in the sector. This retailer has 19 branches nationwide and several brand names linked to it, with 3 divisions: inbound, re-sales and retail. The inbound creates the majority of activity, provides the stock and accounts for 10% of revenue. Re-sales represent 70% of the revenue. Retail reaches the consumer directly and accounts for the remaining 20% of revenue.

Many years' data from multiple outlets tracking all sales operations existed and the company wanted to set up a "scientific platform" to analyse it. In 2014 the data was interrogated differently for the first time. The first problem to arise was variety – inconsistent, incomplete and inaccurate data capture across outlets and delivery channels. Despite being internally generated, data had to be cleaned because of non-standardised data capture processes. There were two product ranges, value and deluxe. Historically the deluxe range had been 'enhanced' by providing a 6 month guarantee. They were surprised to find they lost money on the deluxe and made three times more, per value item, even though it sold more cheaply and it did not have a guarantee. An unexpected pattern in sales was also found: certain models sold for higher prices in certain locations, yet their sales policy was to sell at the location nearest to the previous owner's registered address, regardless of sales value location. Stock days were found to be irrelevant.

These findings meant that the management team reconsidered the positioning of the product range and standardised a number of aspects of their operation hitherto non-standardised.

Data Collection

The operation generates masses of data in a constant stream across all its business units. Data comes from 'in use', 'off use', retail, wholesale, industry regulators and each individual product item. Total sales are approximately 500, 000 units annually. Wanting to create a competitive edge by becoming the most trusted UK provider in its sector the company set up a project to provide a scientific platform to analyse the data collected from the day-to-day business operations. The research questions were brainstormed by the sales team, led by the Heads of Sales and Sales Effectiveness. The aim was to understand the impact of product variety on workload and profitability. The research questions which emerged were:

1. Are products with guarantees more profitable?
2. What links are there between wholesale and retail?
3. What is the link between stock days and product profile?

Hypotheses were developed to run the data mining (DM) verification paradigm as well as enable descriptive analysis:

H1 – product with guarantees are more profitable than those without

H2 – deluxe products are more profitable than standard

The re-sales division had the largest dataset. An initial exploration revealed enough data-sets to answer the research questions.

Data Preparation

A total of 48 variables were found within the sales data. To enable subsequent analysis the variables were defined into: numeric, categorical (string), Boolean (yes/no) and specific codification attributes. Data cleaning had to be carried out due to 3 main inputting problems – inconsistency (eg BLK and BLACK), character transposition and missing values.

Data Analysis

The company adopted the CRISP-DM (Cross Industry Standard Process for Data Mining) approach as the most viable to mine existing data because it is accepted as the 'gold standard' in the data mining domain (Rennolls and Al-Shawabkeh, 2008, Pechenizkiy *et al.*, 2008).

Using this approach it was found that non-guaranteed products' mean contribution to profitability was higher than those with guarantees, therefore **H1** was rejected. Total revenues from deluxe products was found to be 61.5%, therefore **H2** was also rejected. Further exploration of the deluxe category established there was product differentiation within-category. The DM findings identified the most profitable sales.

Findings

The insight that inherently enhanced products were more profitable than those enhanced by guarantee led the management team to reconsider the market positioning of the product range and the input the company provided pre-sales. The descriptive analysis carried out to answer question 2 highlighted the most searched products nationally, by location, which can help decision-making regarding targeted pricing and promotion decisions. The question 3 analysis ascertained there was no particular link between sales and stock days.

The findings from this analysis of all-company data provide a framework for decision-making regarding both strategic and tactical aspects of the business. In addition, it has led to standardisation of data capture to reduce future inconsistency. The company has also recognised that having developed an analysis platform there are many other aspects it can explore operationally. Having established an industry overview, further mining of the data can help it enhance its market leader position. In terms of the framework, this falls into the category of Strategic Support, and containing aspects of Service Design.

6. **Discussion**

Whilst there are limitations within both case studies, they illustrate that there is promise in using BD to uncover previously unavailable insights for sustainable performance.

Both case studies conformed to the Gartner Online (2014) definition of BD: “data assets that require innovative forms of information processing for enhanced insight and decision-making”. Case 1 highlighted the importance of knowing your product and what exactly you wanted to know about your customer and the value of asking interesting and insightful questions. Case 2 highlighted the difficulties encountered with inconsistent data and the importance of reliability and robustness, and showed how to maximise service levels when you use the complete knowledge of long-term process of new sales and used sales over a period of years. These case-specific points relate to the respective organisations, although they can generally be seen to be in keeping with the OECD (2013) suggestion and confirm the importance of developing the correct strategy to advantageously exploit the data (Opresnik and Taisch, 2015). They are also problems consistent with the core concerns of Operations Management which scholars have studied continuously for many years (Ackoff, 1989, Kranzberg, 1986, Bell, 1999, Saaty, 2008). It is this aspect which has led to observations that BD is another step in data-driven improvement, following in the footsteps of SPC and then 6 Sigma (Madsen and Stenheim, 2013, Näslund, 2008)

The analysis of the competitive landscape relating to BD vendors enables a more granular definition of “*insight and decision-making*”. Each of the two case studies demonstrated different elements of the BD framework in use. The framework could be used to guide the business in Case 1 to extend their use of personalisation into the area of prediction of product bundles that will be desired by customers, and thence to improved service design and responsiveness. Equally the business in Case 2 was described as beginning to use its BD driven insights to redesign its portfolio of services and to apply these in a targeted way to product categories, an extension of BD use predicted by the framework.

Case 1 wanted to enhance the customer experience. Case 2 wanted to become the most trusted UK provider in its market. Both companies used internal data about committed customers to create competitive advantage. Unlike GFT, where individuals other than those suffering flu would make searches about the topic, in this case the stream of data was a complete picture of the companies’ customers and their behaviour. The use of BD techniques meant that is was unnecessary to draw samples of behaviour and try to use this to predict individual responses. By using BD techniques both companies were able to maximise their management decisions, based on a complete picture of the behaviour of their total customer base. This is a distinctly better method of analysis then simply looking for trends in the whole population and trying to relate it to their business, as was being done with GFT. It is also a true application of BD, whereas the GFT case, despite collating social media and other externally-generated data was actually using a population sample.

Sampling and hypothesis-building is traditional data analysis, not BD, despite the constant stream of largely unstructured data created from multiple sources and in a range of formats which causes the BD analytics challenge (Bisel *et al.*, 2014, Chen *et al.*, 2012, Tinati *et al.*, 2014).

Case 2's data, whilst not quite unstructured, lacked a standardised company format. The same products were differently captured on the system by each branch, creating a cleaning project first, then a project to systematically identically capture data throughout the organisation, regardless of product or channel.

The key aspect for both cases is the importance of initial inspiration by individual employees. Neither case would have happened had a member of staff not wondered about a particular problem, just as with Wald in 1943. Technology enables the speedier realisation of an answer but it cannot show the direction to choose for exploration. So, far from being helpful, BD, as stated in the literature review, has enormous potential for misinterpretation and misdirection, of which numerous examples exist (Lazer *et al.*, 2014, Ilic *et al.*, 2010, Hazen *et al.*, 2014, Duhigg, 2012). Whilst the framework developed from this research can help, it is specific people skills which are required to outline the initial enquiry. There is a clear need, as shown by the case studies, to have employees with a level of curiosity which can lead to new streams of exploration. Relying on algorithms and system tools is unlikely to present new insights. The key challenge therefore is to identify the new skills people need in order to be able to maximise the potential BD offers and underlines the relevance of the GCHQ development (Philipson, 2014).

These 2 empirical studies are the first step in validating the BD framework, which can help organisations in two key ways. Firstly, it can help guide the nature of the questions required based on the outcome sought. Secondly, this helps shape an understanding of the people skills required for both interpreting data-sets and developing relevant, impactful insight. The white goods retailer highlighted the importance of knowing your product and what exactly you wanted to know about your customer. The mixed retailer data analysis highlighted the difficulties encountered with inconsistent data and the importance of reliability and robustness. More examples and tests are needed.

7. **Conclusion**

This paper has explored the challenges and opportunities of BD. The research shows that despite the hype BD solutions are being used to manage and improve operational processes. Irrespective of the longevity of the term and the novelty of the techniques, the data available to organisations will continue to grow and technology will be increasingly utilised to harness the value locked away in that data. The challenge to harnessing the full power of the technology is in formulating questions and new perspectives from which to examine problems or perceived problems and the answers to these problems. This then enables organisations to exploit and leverage the power of the technology to create sustainable competitive advantage. The framework provides a conceptual model to target desired outcomes and the questions to get there.

The collection, processing and utilisation of data to inform decision-making and optimise processes are defining features of operations management and operations research. Operations are already being controlled and automated using advanced technologies and the availability of more data and sophisticated analytic tools will be exploited in the future. Yet limitless data does not guarantee better operational performance. Operations managers and researchers need to learn how to use new tools and develop the cognitive skills necessary to generate the knowledge and wisdom they will require to manage operations within an increasingly connected, complex and rapidly changing world.

BD is identified in three main areas: administrative (from public bodies), Social media and private sector data. A number of problems emerge, which are also opportunities for further research. There are access and reliability problems. Understanding the utility of the data as research resources is paramount. Given the variety of sources, data reuse, linkage and sharing may be difficult. How to overcome these issues, and uncover if it is possible to do so reliably, is a fruitful avenue for future researchers.

There are problems with how people think, react and respond to things at any given time. For instance the way speed is seen affects how data is perceived. Speed may mean better, more informed decisions can be made by individuals for themselves and for their organisations. Or it may mean greater confusion due to rapidly changing scenarios having to be grappled with. Operations Management capacity issues now need to also incorporate how much data people can cope with to optimise what is done and how it is done. Behavioural Operations has a key role to play in researching whether behaviour is influenced by the data presented or whether behaviour affects the data. Changes in how customers are approached and dealt with also need to be considered. Just as organisations are impacted by BD, so are customers. They have more information at their fingertips than ever before. Speed becomes a factor for them too, as does knowledge. Organisations need to understand what customers do with the knowledge they have, how they exercise their decision-making choices and if/how organisations can influence that.

Staff requirements also need to be rethought, and research into this is paramount. There needs to be a step change in competences for some of the tasks expected to be done as a matter of course once the world of BD is embraced. New data science skill-sets are required, because current practices, without insightful and curious employees, mean more data is being churned more rapidly but not necessarily for greater advantage. Indeed, the GFT example shows how it can impede true knowledge by being inherently flawed. The key challenge therefore is to identify the new skills people need in order to be able to maximise the potential BD offers, as GCHQ have already done, although much more work is needed.

Finally, data curation needs have to be rethought. As yet there is no evidence how this is going to be tackled in a sustainable way, or indeed if it will be.

The framework proposed in this study offers a potentially useful starting point for understanding where and how organisations can leverage business value from big data. The case studies demonstrate two of the categories in use. Further research is required to fully validate the remaining categories within the framework and to ascertain its utility. Future organisational success, if not survival, may well be predicated on insightful application of systems-generated, people-enhanced information.

References

Ackoff, R. (1989) "From Data to Wisdom". *Journal of Applied Systems Analysis,* Vol. 16 No.**,** pp 3–9.

Barton, D. and Court, D. (2012) "Making Advanced Analytics Work for You". *Harvard Business Review,* Vol. 90 No. 10**,** pp 78-83.

Bell, P. (2013) "Creating Competitive Advantage Using Big Data". *Ivey Business Journal,* Vol. 77 No. 3**,** pp 4-8.

Bell, P. C. (1999) *Management Science/Operations Research: A Strategic Perspective.* South-Western College Publishing Boston, MA

Bisel, R. S., Barge, J. K., Dougherty, D. S., Lucas, K. and Tracy, S. J. (2014) "A Round-Table Discussion of "Big" Data in Qualitative Organizational Communication Research". *Management Communication Quarterly,* Vol. 28 No. 4**,** pp 625-649.

Bollen, J., Mao, H. and Zeng, X. (2011) "Twitter Mood Predicts the Stock Market". *Journal of Computational Science,* Vol. 2 No. 1**,** pp 1-8.

Brown, B., Chui, M. and Manyika, J. (2011) "Are You Ready for the Era of 'Big Data'?". *McKinsey Quarterly,* Vol. 4 No.**,** pp 24-35.

Browser-update.org (2015) *Homepage Statement.* Available from: https://www.linkedin.com/company/6202954?trk=vsrp_companies_cluster_name&trkInfo=VSRPsearchId%3A32485231427721498574%2CVSRPtargetId%3A6202954%2CVSRPcmpt%3Acompanies_cluster (Accessed 30th March 2015).

Bughin, J., Livingston, J. and Marwaha, S. (2011) "Seizing the Potential of 'Big Data'". *McKinsey Quarterly,* Vol. 4 No.**,** pp 103-109.

Carvajal, R. (1992) "Operations Research (or), Management Science (Ms), Systems Science, and Russell Ackoff: The Development of Two Paradigms". *Systems practice,* Vol. 5 No. 3**,** pp 291-318.

Chae, B. (2015) "Insights from Hashtag #Supplychain and Twitter Analytics: Considering Twitter and Twitter Data for Supply Chain Practice and Research". *International Journal of Production Economics,* Vol. 165 No.**,** pp 247-259.

Chen, H., Chiang, R. H. L. and Storey, V. C. (2012) "Business Intelligence and Analytics: From Big Data to Big Impact". *MIS Quarterly,* Vol. 36 No. 4**,** pp 1165-1188.

Chongwatpol, J. and Chan, H. K. (2015) "Prognostic Analysis of Defects in Manufacturing". *Industrial Management & Data Systems,* Vol. 115 No. 1**,** pp 64-87.

Cox, M. and Ellsworth, D. (1997) Application-Controlled Demand Paging for out-of-Core Visualization. In: *Proceedings of the 8th conference on Visualization '97.* Phoenix, Arizona, USA: IEEE Computer Society Press, pp. 235-243.

Croxall, B. (2014) *The Red Herring of Big Data.* Brian Croxall web post. Available from: http://www.briancroxall.net/2014/03/07/the-red-herring-of-big-data/ (Accessed 30th March 2015).

Demirkan, H. and Delen, D. (2013) "Leveraging the Capabilities of Service-Oriented Decision Support Systems: Putting Analytics and Big Data in Cloud". *Decision Support Systems,* Vol. 55 No. 1**,** pp 412-421.

Diebold, F. X. (2012) "A Personal Perspective on the Origin(S) and Development of 'Big Data': The Phenomenon, the Term, and and the Discipline". *PIER Working Paper,* Vol. Second Version (26 November, 2012) No. 13-003**,** pp.

Dubey, R., Gunasekaran, A., Childe, S., Wamba, S. and Papadopoulos, T. (2015) "The Impact of Big Data on World-Class Sustainable Manufacturing". *The International Journal of Advanced Manufacturing Technology,* Vol. No.**,** pp 1-15.

Duhigg, C. (2012) How Companies Learn Your Secrets. *New York Times Magazine.* NY Times.

Eisenhardt, K., M. (2002) Building Theories from Case Study Research. In: Huberman, A. M. andMiles, M. B. (Eds.) *The Qualitative Researcher's Companion.* Thousand Oaks; London: Sage Publications, pp. 5-35.

Esener, S. C., Kryder, M. H., Doyle, W. D., Keshner, M., Mansuripur, M. and Thompson, D. A. (1999) *The Future of Data Storage Technologies.* PB99-144214. Baltimore, Maryland International Technology Research Institute.

Fawcett, S. E. and Waller, M. A. (2014) "Supply Chain Game Changers—Mega, Nano, and Virtual Trends—and Forces That Impede Supply Chain Design (I.E., Building a Winning Team)". *Journal of Business Logistics,* Vol. 35 No. 3**,** pp 157-164.

Fisher, D., DeLine, R., Czerwinski, M. and Drucker, S. (2012) "Interactions with Big Data Analytics". *interactions,* Vol. 19 No. 3**,** pp 50-59.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015) "How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study". *International Journal of Production Economics,* Vol. 165 No. July**,** pp 234-246.

Gandomi, A. and Haider, M. (2015) "Beyond the Hype: Big Data Concepts, Methods, and Analytics". *International Journal of Information Management,* Vol. 35 No. 2**,** pp 137-144.

Gartner (2013) *Emerging Technologies Hype Cycle.* Gartner Research and Advisory. Available from: http://www.gartner.com/newsroom/id/2575515 (Accessed 14th September 2015).

Gartner (2014) *Gartner Says Annual Smartphone Sales Surpassed Sales of Feature Phones for the First Time in 2013.*

Gartner (2015) *Emerging Technologies Hype Cycle.* Gartner Research and Advisory. Available from: http://www.gartner.com/newsroom/id/3114217 (Accessed 14th September).

George, G., Haas, M. R. and Pentland, A. (2014) "Big Data and Management". *Academy of Management Journal,* Vol. 57 No. 2**,** pp 321-326.

Glaser, B. G. and Strauss, A. L. (1968) *The Discovery of Grounded Theory : Strategies for Qualitative Research.* (Observations) Weidenfeld & Nicolson, London

Haas, M. R., Criscuolo, P. and George, G. (2014) "Which Problems to Solve? Attention Allocation and Online Knowledge Sharing in Organizations". *Academy of Management Journal,* Vol. No.**,** pp.

Hair, J. F. H., Money, A. H., Samouel, P. and Page, M. (2007) *Research Methods for Business.* Wiley, Chichester

Hazen, B. T., Boone, C. A., Ezell, J. D. and Jones-Farmer, L. A. (2014) "Data Quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An Introduction to the Problem and Suggestions for Research and Applications". *International Journal of Production Economics,* Vol. 154 No. 0**,** pp 72-80.

Housley, W., Procter, R., Edwards, A., Burnap, P., Williams, M., Sloan, L., Rana, O., Morgan, J., Voss, A. and Greenhill, A. (2014) "Big and Broad Social Data and the Sociological Imagination: A Collaborative Response". *Big Data & Society,* Vol. 1 No. 2**,** pp.

Hruska, J. (2013) *Intel's Former Chief Architect: Moore's Law Will Be Dead within a Decade.* ExtremeTech. Available from: http://www.extremetech.com/computing/165331-intels-former-chief-architect-moores-law-will-be-dead-within-a-decade (Accessed 30th March 2015).

Huang, Y.-Y. and Handfield, R. B. (2015) "Measuring the Benefits of Erp on Supply Management Maturity Model: A "Big Data" Method". *International Journal of Operations & Production Management,* Vol. 35 No. 1**,** pp 2-25.

IDC (2012) *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East.* Online:

Ilic, A., Grössbauer, A., Michahelles, F. and Fleisch, E. (2010) "Understanding Data Volume Problems of Rfid ‐ Enabled Supply Chains". *Business Process Management Journal,* Vol. 16 No. 6**,** pp 904-916.

Jifa, G. and Lingling, Z. (2014) "Data, Dikw, Big Data and Data Science". *Procedia Computer Science,* Vol. 31 No. 0**,** pp 814-821.

Kelly, J. (2014) *Big Data Vendor Revenue and Market Forecast 2013-2017.* Wikibon. Available from: http://wikibon.org/wiki/v/Big_Data_Vendor_Revenue_and_Market_Forecast_2013-2017 (Accessed 19th March 2015).

Kranzberg, M. (1986) "Technology and History: "Kranzberg's Laws"". *Technology and Culture,* Vol. 27 No. 3**,** pp 544-560.

Laney, D. (2001) 3d Data Management: Controlling Data Volume, Velocity and Variety. *Application Delivery Strategies*, 4. Available from http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf (Accessed 19th March 2015)

Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014) "The Parable of Google Flu: Traps in Big

Data Analysis". *Science,* Vol. 343 No. 6176, pp 1203-1205.

Lee, H. and Özer, Ö. (2007) "Unlocking the Value of Rfid". *Production and Operations Management,* Vol. 16 No. 1, pp 40-64.

Li, X., Song, J. and Huang, B. (2015) "A Scientific Workflow Management System Architecture and Its Scheduling Based on Cloud Service Platform for Manufacturing Big Data Analytics". *The International Journal of Advanced Manufacturing Technology,* Vol. No., pp 1-13.

Liew, P. (2007) "Understanding Data, Information, Knowledge and Their Inter-Relationships". *Journal of Knowledge Management Practice.,* Vol. 8 No. 2, pp.

Lucke, D., Constantinescu, C. and Westkämper, E. (2008) Smart Factory - a Step Towards the Next Generation of Manufacturing. In: Mitsuishi, M., et al. (Eds.) *Manufacturing Systems and Technologies for the New Frontier.* Springer London, pp. 115-118.

Lury, C. (2013) "Focus: Big Data, Little Questions? ". *Discover Society,* Vol. 1 No. October, pp.

Madsen, D. Ø. and Stenheim, T. (2013) "Doing Research on "Management Fashions" : Methodological Challenges and Opportunities". *Problems and perspectives in management,* Vol. 11 No. 4, pp 68-76.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Hung Byers, A. (2011) *Big Data: The Next Frontier for Innovation, Competition and Producivity.* McKinsey Global Institute.

McAfee, A. and Brynjolfsson, E. (2012) "Big Data: The Management Revolution. (Cover Story)". *Harvard Business Review,* Vol. 90 No. 10, pp 60-68.

Meyer, G. G., Buijs, P., Szirbik, N. B. and Wortmann, J. C. (2014) "Intelligent Products for Enhancing the Utilization of Tracking Technology in Transportation". *International Journal of Operations & Production Management,* Vol. 34 No. 4, pp 422-446.

Miller, S. (2014) "Collaborative Approaches Needed to Close the Big Data Skills Gap". *Journal of Organization Design,* Vol. 3 No. 1, pp 26-30.

Moore, G. E. (1965) "Cramming More Components onto Integrated Circuits". *Electronics,* Vol. 38 No. 8, pp 114-117.

Näslund, D. (2008) "Lean, Six Sigma and Lean Sigma: Fads or Real Process Improvement Methods?". *Business Process Management Journal,* Vol. 14 No. 3, pp 269-287.

O'Leary, D. E. (2013) "'Big Data', the 'Internet of Things' and the 'Internet of Signs'". *Intelligent Systems in Accounting, Finance and Management,* Vol. 20 No. 1, pp 53-65.

OECD (2013) *Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data".* OECD Digital Economy Papers, No 222: OECD Publishing.

Opresnik, D. and Taisch, M. (2015) "The Value of Big Data in Servitization". *International Journal of Production Economics,* Vol. 165 No. July, pp 174-184.

Pechenizkiy, M., Puuronen, S. and Tsymbal, A. (2008) "Towards More Relevance-Oriented Data Mining Research". *Intelligent Data Analysis,* Vol. 12 No. 2, pp 237-249.

Philipson, A. (2014) *Gchq Employs More Than 100 Dyslexic and Dyspraxic Spies.* [http://www.telegraph.co.uk/education/educationnews/11111584/GCHQ-employs-more-than-100-dyslexic-and-dyspraxic-spies.html] (Accessed 19th January 2015).

Piccoli, G. and Pigni, F. (2013) "Harvesting External Data: The Potential of Digital Data Streams". *MIS Quarterly Executive,* Vol. 12 No. 1, pp 53-64.

Porter, M. E. and Heppelmann, J. E. (2014) "How Smart, Connected Products Are Transforming Competition". *Harvard Business Review,* Vol. 92 No. 11**,** pp 64-88.

Rennolls, K. and Al-Shawabkeh, A. (2008) "Formal Structures for Data Mining, Knowledge Discovery and Communication in a Knowledge Management Environment". *Intelligent Data Analysis,* Vol. 12 No. 2**,** pp 147-163.

Rowley, J. (2007) "The Wisdom Hierarchy: Representations of the Dikw Hierarchy". *Journal of Information Science,* Vol. 33 No. 2**,** pp 163-180.

Saaty, T. L. (2008) "Decision Making with the Analytic Hierarchy Process". *International Journal of Services Sciences,* Vol. 1 No. 1**,** pp 83-98.

Sætre, A. S., Sørnes, J.-O., Browning, L. D. and Stephens, K. K. (2003) Organizational Members' Enactment of Organizational Environments and Media Use: A Study of Ict Practicesin Norway and the United States. In: *Informing Science and Information Technology Education Joint Conference.* Pori, Finland:

Segerberg, A. and Bennett, W. L. (2011) "Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of Two Climate Change Protests". *The Communication Review,* Vol. 14 No. 3**,** pp 197-215.

Shedroff, N. (1994) "Information Interaction Design a Unified Field Theory of Design". Vol. No.**,** pp.

Tinati, R., Halford, S. and Carr, L. (2014) "The Promise of Big Data: New Methods for Sociological Analysis". *Sociology,* Vol. 48 No. 4**,** pp 663-681.

Wald, A. (1980) *A Reprint of 'a Method of Estimating Plane Vulnerability Based on Damage of Survivors.* CRC 432. CENTER FOR NAVAL ANALYSES ALEXANDRIA VA: OPERATIONS EVALUATION GROUP.

Wieland, A., Handfield, R. B. and Durach, C. F. (2014) Mapping the Landscape of Emerging Research Topics in Supply Chain Management. In: *Proceedings of 21st EurOMA Conference.* Palermo, Sicily: European Operations Management Association, 2014 p. 9.

Wise, G. L. (1974) "Differential Pricing and Treatment by New-Car Salesmen: The Effect of the Prospect's Race, Sex, and Dress". *Journal of Business,* Vol. 47 No. 2**,** pp 218-230.

Zelbst, P. J., Green, K. W., Sower, V. E. and Reyes, P. M. (2012) "Impact of Rfid on Manufacturing Effectiveness and Efficiency". *International Journal of Operations & Production Management,* Vol. 32 No. 3**,** pp 329-350.

Zhong, R., Y, Lan, S., Xu, C., Dai, Q. and Huang, G. (2015) "Visualization of Rfid-Enabled Shopfloor Logistics Big Data in Cloud Manufacturing". *The International Journal of Advanced Manufacturing Technology,* Vol. No.**,** pp 1-12.